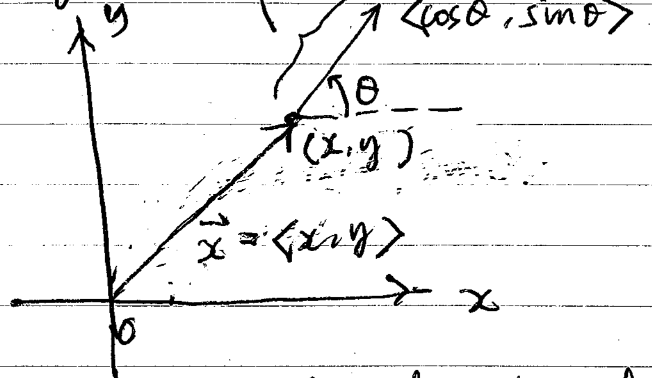
Directional Derivatives of $f(x,y)$

This is a generalization of $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ for a function $f(x,y)$. Let now $\vec{u} = \langle \cos\theta, \sin\theta \rangle$ be a direction vector in the xy plane (recall, a direction vector $\vec{u}$ is a vector pointing in a certain direction of length 1).



We define the directional derivative of f in the direction $\vec{u}$ as follows. We use vector notation with $\vec{x} = \langle x, y \rangle$ and $\vec{x} + h\vec{u} = \langle x + h\cos\theta, y + h\sin\theta \rangle$ and define

$$\frac{\partial f(\vec{x})}{\partial \vec{u}} = \lim_{h\to 0} \frac{f(\vec{x} + h\vec{u}) - f(\vec{x})}{h}$$

$$= \lim_{h \to 0} \frac{f(x + h\cos\theta, y + h\sin\theta) - f(x,y)}{h}$$

We construct the auxillary

$$g(t) = f(x + t\cos\theta, y + t\sin\theta)$$

then $g'(t) = \frac{\partial f}{\partial x}(x + t\cos\theta, y + t\sin\theta)\cos\theta + \frac{\partial f}{\partial y}(x + t\cos\theta, y + t\sin\theta)\sin\theta$

In particular, $g'(0) = \frac{\partial f}{\partial x}(x,y)\cos\theta + \frac{\partial f}{\partial y}(x,y)\sin\theta$

But $g'(0) = \lim_{h\to 0} \frac{g(h) - g(0)}{h}$ also,

thus $$\lim_{h\to 0} \frac{f(x + h\cos\theta, y + h\sin\theta) - f(x,y)}{h} = \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta$$

Remark: Note that $\frac{\partial f}{\partial \vec{i}}(x,y) = \frac{\partial f}{\partial x}(x,y)$, $\frac{\partial f}{\partial \vec{j}}(x,y) = \frac{\partial f}{\partial y}(x,y)$

In general, suppose we have a function of n variables $f(x_1,\cdots,x_n)$, set $\vec{x}=\langle x_1,\cdots,x_n\rangle$ and let $\vec{u}=\langle u_1,\cdots,u_n\rangle$, $|\vec{u}|=1$ be any direction vector, we define

$$\frac{\partial f}{\partial \vec{u}}(\vec{x}) = \lim_{h\to 0}\frac{f(\vec{x}+h\vec{u})-f(\vec{x})}{h}$$

to be the directional derivative of $f(\vec{x})$ in the direction $\vec{u}$.

Definition: Given $f(\vec{x})=f(x_1,\cdots,x_n)$, $\vec{x}=(x_1,\cdots,x_n)$, we define the gradient vector of f by

$$\nabla f(\vec{x}) = \left\langle \frac{\partial f}{\partial x_1}(x_1,\cdots,x_n), \frac{\partial f}{\partial x_2}(x_1\cdots x_n), \cdots, \frac{\partial f}{\partial x_n}(x_1,\cdots,x_n)\right\rangle$$

Then along any direction $\vec{u}$, we have

$$\boxed{\frac{\partial f(\vec{x})}{\partial \vec{u}} = \nabla f(\vec{x})\cdot\vec{u} = \frac{\partial f}{\partial x_1}(x_1\cdots x_n)u_1 + \cdots + \frac{\partial f}{\partial x_n}(x_1\cdots x_n)u_n}$$

Remark: $\frac{\partial f}{\partial \vec{u}}$ is the rate of change of f in the direction of $\vec{u}$.

Ex. Given $f(x,y,z)=x^2+y^2+z^2-2xy+2xz+yz$, find the directional derivative of f at the point $P(1,0,2)$ in the direction from P to $Q(2,3,3)$

Solution:

$$\nabla f(x,y,z) = \langle 2x-2y+2z,\ 2y-2x+z,\ 2z+2x+y\rangle$$

$$\therefore \nabla f(1,0,2) = \langle 2+4,\ -2+2,\ 4+2\rangle = \langle 6,0,6\rangle$$

and the direction vector in the direction of $\vec{PQ}=\langle 1,3,1\rangle$ is given by $\vec{u}=\frac{\langle 1,3,1\rangle}{\sqrt{1+9+1}} = \left\langle \frac{1}{\sqrt{11}}, \frac{3}{\sqrt{11}}, \frac{1}{\sqrt{11}}\right\rangle$

$$\therefore \frac{\partial f}{\partial \vec{u}}(1,0,2) = \nabla f(1,0,2)\cdot\vec{u} = \langle 6,0,6\rangle\cdot\left\langle \frac{1}{\sqrt{11}}, \frac{3}{\sqrt{11}}, \frac{1}{\sqrt{11}}\right\rangle = \frac{12}{\sqrt{11}}$$

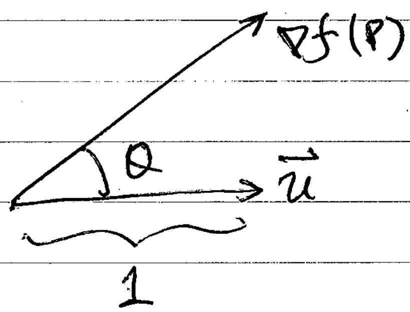
Direction in which f changes most rapidly and the Direction in which the rate of change ~~of~~ of f is a minimum.

Let $P(x_0, y_0, z_0)$ be any point belonging to the domain of f and let $\vec{u}$ be any direction, then

$$\frac{\partial f}{\partial \vec{u}}(P) = \nabla f(P) \cdot \vec{u}$$

By dot product's property, $\frac{\partial f}{\partial u}(P) = |\nabla f(P)||\vec{u}|\cos\theta = |\nabla f(P)|\cos\theta$

where θ is the $\measuredangle$ between $\nabla f(P)$ and $\vec{u}$



We observe that, $\frac{\partial f}{\partial u}(P)$ attains its maximum value of $|\nabla f(P)|$ when $\theta = 0$, i.e. when $\vec{u}$ is in the same direction of $\nabla f(P)$.

Similarly, the rate of change of f i.e. $\frac{\partial f}{\partial u}(P)$ attains its minimum value of $-|\nabla f(P)|$ when $\vec{u}$ is in the opposite direction of $-\nabla f(P)$.

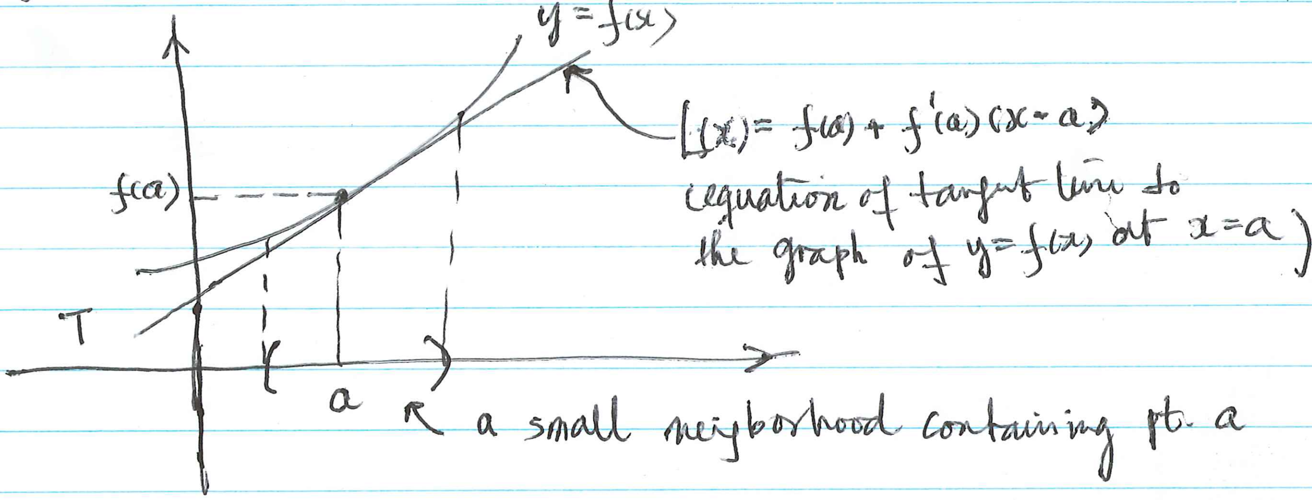
Ex Suppose the temperature T (in °C) at the pt. (x, y, z) in space is given by $T(x,y,z) = x^3 - xy^2 - z$, in which direction does the temperature increase most rapidly at $P(1,1,0)$? What is the maximal rate of change at $P(1,1,0)$?

<u>Solution:</u> $\nabla T(x,y,z) = \langle 3x^2 - y^2, -2xy, -1 \rangle$

at $P(1,1,0)$, maximum rate of change of T is in the direction of $\nabla T(1,1,0) = \langle 2, -2, -1 \rangle$ or $\vec{u} = \langle \frac{2}{3}, \frac{-2}{3}, \frac{-1}{3} \rangle$ with max rate of change $= |\nabla T(1,1,0)| = 3$ //

Differentials and Linear Approximation

Revisiting the function of one variable case,



In a small neighborhood containing the pt. a, we could use the tangent line $y = L(x) = f(a) + f'(a)(x-a)$ as an approximation for $y = f(x)$.

Thus,

$$f(x) \simeq f(a) + f'(a)(x-a)$$

Suppose now we go from $x = a$ to $x = a + h$, we have by approximation,

$$f(a+h) \simeq f(a) + f'(a)h$$

$$\Rightarrow \quad f(a+h) - f(a) \simeq f'(a)h$$

Returning to a general point x, we have

$$\boxed{f(x+h) - f(x) \simeq f'(x)h \quad \text{for small } h.} \quad \text{-----} \; (*)$$

We define the differentials dx and dy (or df) as,

$dx = \Delta x$ or h which is a small change in x.

$dy = f'(x)\,dx$ or $f'(x)\Delta x$

Finally, to distinguish it from dy, we define

$\Delta y = f(x+h) - f(x)$ or $f(x+\Delta x) - f(x)$

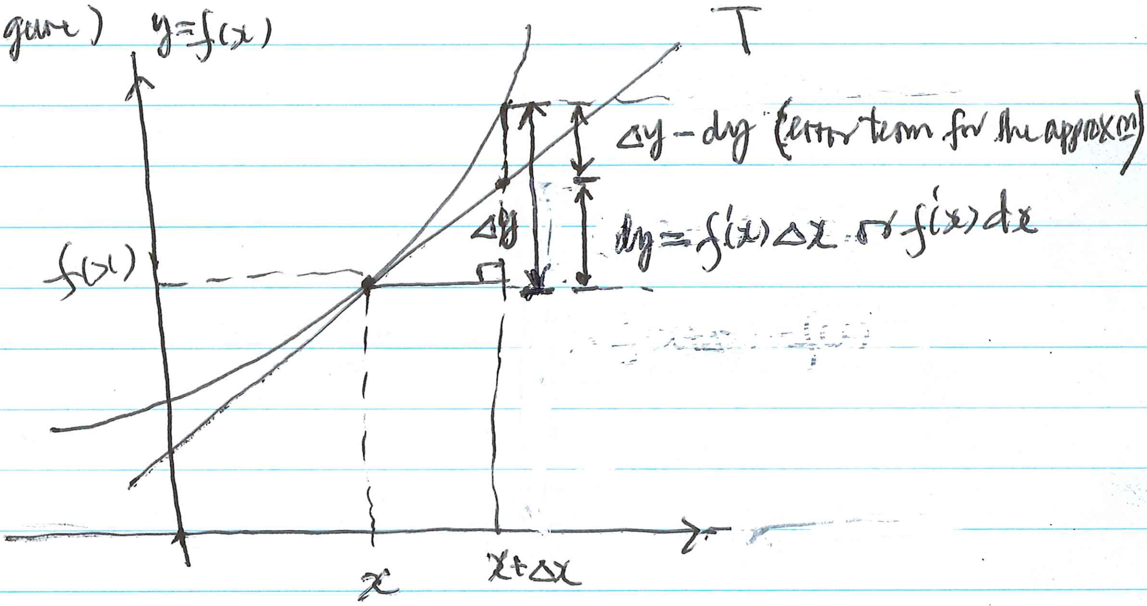
which is the exact change in f when we go from x to $x + \Delta x$

In view of (*),

$$\underbrace{f(x+\Delta x)-f(x)}_{\Delta y} \simeq f'(x)\Delta x$$

$$\Delta y \simeq dy.$$

Hence, while for the independent variable x, we define the differential of x, dx the same as Δx, for the dependent variable y, dy is not the same as Δy, the exact change in $f(x)$ as we go from x to $x+\Delta x$ (see figure) $y=f(x)$



Generalizing to the case $z=f(x,y)$, we could use the tangent plane

$$z = L(x,y) = f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial x}(a,b)(y-a)$$

as an approximation for the surface $z=f(x)$ nearby the pt. $(a,b,f(a,b))$ on the surface. Arguing in exactly the same way, we could define

$dx = \Delta x$ (change in x)

$dy = \Delta y$ (" " y)

dz or $df = \dfrac{\partial f}{\partial x}(x,y)\,dx + \dfrac{\partial f}{\partial y}(x,y)\,dy$

Δz or $\Delta f = f(x+\Delta x, y+\Delta y) - f(x,y)$

$dz \simeq \Delta z$ or Δf for small Δx & small Δy.

Ex. Find an approximate value for $\sqrt{(3\cdot03)^2+(3\cdot99)^2}$

Solution:

We construct $f(x,y)=\sqrt{x^2+y^2}$ and observe that $(3\cdot03, 3\cdot99)$ is close to the pt. $(3,4)$ with $f(3,4)=5$.

From $(x,y)=(3,4)$ to $(x+\Delta x, y+\Delta y)=(3\cdot03, 3\cdot99)$,

$dx=0\cdot03$, $dy=-0\cdot01$

The corresponding change in f, Δf is approximated by

$$df=\frac{\partial f}{\partial x}(3,4)dx+\frac{\partial f}{\partial y}(3,4)dy$$

$$\frac{\partial f}{\partial x}=\frac{x}{\sqrt{x^2+y^2}}, \quad \frac{\partial f}{\partial y}=\frac{y}{\sqrt{x^2+y^2}}$$

$$\therefore df=\left(\frac{3}{5}\right)(0\cdot03)+\left(\frac{4}{5}\right)(-0\cdot01)=\frac{0\cdot09-0\cdot04}{5}=0\cdot01$$

Thus $f(3\cdot03, 3\cdot99)\simeq f(3,4)+0\cdot01=5\cdot01$ //

Generalizing to the function of n variable case,

$$w=f(x_1,\cdots,x_n)$$

we define $dx_1=\Delta x_1, \cdots, dx_n=\Delta x_n$

as the change in x_1, ---- change in x_n respectively.

and define

$$dw \text{ or } df=\frac{\partial f}{\partial x_1}(x_1,\cdots,x_n)dx_1+\cdots+\frac{\partial f}{\partial x_n}(x_1\cdots x_n)dx_n$$

we have $\quad df\simeq\Delta f$

where $\Delta f=f(x_1+\Delta x_1,\cdots,x_n+\Delta x_n)-f(x_1,\cdots,x_n)$

is the exact change in f when we go from $(x_1\cdots x_n)$ to $(x_1+\Delta x_1,\cdots,x_n+\Delta x_n)$.

Remark: In vector notation, we could set

$\vec{x} = \langle x_1, \cdots, x_n \rangle$, $\vec{a} = \langle a_1, \cdots, a_n \rangle$ be a pt in the domain of $f(\vec{x})$

we have the linear approximation of $f(\vec{x})$ at $\vec{a}$,

$$\boxed{w = L(\vec{x}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a})} \quad \text{where } \nabla f = \left\langle \frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_n} \right\rangle$$

i.e.

$$L(x_1, \cdots, x_n) = f(a_1, \cdots, a_n) + \frac{\partial f}{\partial x_1}(a_1, \cdots, a_n)(x_1 - a_1) + \cdots + \frac{\partial f}{\partial x_n}(a_1, \cdots, a_n)(x_n - a_n)$$

which is the equation of the tangent plane to the surface $w = f(x_1 \cdots x_n)$

at the point $\vec{a} = \langle a_1, a_2, \cdots, a_n \rangle$.

we define

$$\boxed{d\vec{x} \text{ or } \Delta\vec{x} = \langle dx_1, \cdots, dx_n \rangle}$$

Then $$\boxed{dw \text{ or } df = \nabla f(\vec{x}) \cdot d\vec{x}}$$

which is an approximation for $\boxed{\Delta f = f(\vec{x} + \Delta\vec{x}) - f(\vec{x})}$, the exact change in f from $\vec{x}$ to $\vec{x} + \Delta x$.

Concept of Differentiability for multi-variable Functions:

Natural Question:

As we go from $f(x)$ to $f(x_1, \cdots, x_n)$, since we now have n variables or n dimensions to change in the domain of f, we have naturally introduced rates of change of f along the n dimensions. This amounts to the n partial derivatives:

$$\forall\, 1 \le i \le n, \quad \frac{\partial f}{\partial x_i}(x_1, \cdots, x_i, \cdots x_n) = \lim_{h \to 0} \frac{f(x_1, \cdots, x_i + h, \cdots x_n) - f(x_1, \cdots, x_i, \cdots x_n)}{h}$$

Is this enough?

Later on, we introduced the concept of directional derivative,

$$\frac{\partial f(\vec{x})}{\partial \vec{u}} = \lim_{h\to 0} \frac{f(\vec{x}+h\vec{u})-f(\vec{x})}{h} \quad , \vec{x}=\langle x_1,\cdots,x_n\rangle$$

so that f is allowed to change not just along the standard n dimensions $\vec{u_1}=\langle 1,0\cdots 0\rangle$, $\vec{u_2}=\langle 0,1,0\cdots 0\rangle \cdots \vec{u_n}=\langle 0,0\cdots 0,1\rangle$ but along any directions $\vec{u}$ with $|\vec{u}|=1$.

But this is still far from being a satisfactory theory because of the following reasons.

(A) From the geometric point of view, differentiability is somewhat equivalent to smoothness in the 1-D setting. If $f(x)$ is differentiable at a pt. a, i.e. $f'(a)$ exists, then $y=f(x)$ is smooth at the pt. $(a, f(a))$ with ~~the~~ the existence of a tangent line touching the curve/graph at that pt. Least to say, $f'(a)$ exists $\Rightarrow f$ must be continuous at the pt. a i.e. differentiability $\Rightarrow$ continuity. The serious problem is, for multi-variable functions, this is no longer true, for partial differentiability doesn't necessarily imply differentiability. We could take a look at the following example.

$$f(x,y)=\begin{cases} \dfrac{xy}{x^2+y^2} & \text{for } (x,y)\neq(0,0) \\ 0 & \text{for } (x,y)=(0,0)\end{cases}$$

We observe that

(i) $\dfrac{\partial f}{\partial x}(0,0) = \lim\limits_{h\to 0}\dfrac{f(h,0)-f(0,0)}{h} = 0$

$\dfrac{\partial f}{\partial y}(0,0) = \lim\limits_{h\to 0}\dfrac{f(0,h)-f(0,0)}{h} = 0$

i.e. f is both partially differentiable with respect to x & to y at $(0,0)$.

(ii) However, in this case $\lim\limits_{(x,y)\to(0,0)} \dfrac{xy}{x^2+y^2}$ does not even exist.

Hence, f has a non-removable discontinuity at $(0,0)$. This counter example has reflected partial differentiability is not an adequate concept or complete concept for differentiability.

(B) From an analytic point of view, in any direction $\vec{u}$, when we define the directional derivative

$$\frac{\partial f}{\partial \vec{u}}(\vec{x}) = \lim_{h \to 0} \frac{f(\vec{x}+h\vec{u}) - f(\vec{x})}{h}$$

we are confined to approach the point $x = (x_1, \cdots, x_n)$ only along a specific direction $\vec{u}$. Analogous to the function of one variable case where we have

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h},$$

we wish to have something similar in the n-D setting, such as

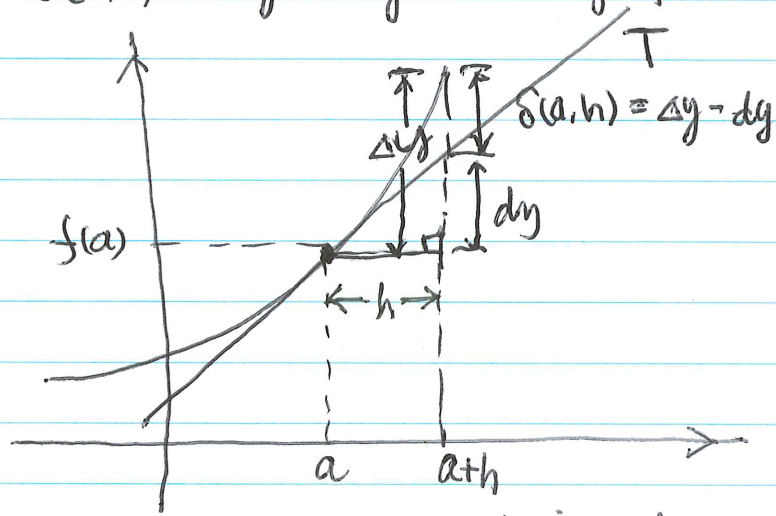
$$Df(\vec{x}) = \lim_{\vec{h} \to \vec{0}} \frac{f(\vec{x}+\vec{h}) - f(\vec{x})}{\vec{h}}$$

where $\vec{x}+\vec{h}$ could approach $\vec{x}$ in any random way without being confined to be along a certain fixed direction. If the above formulation makes sense, we could then call $Df(x)$, the <u>Total Derivative</u> of f in contrast to its partial derivatives or its directional derivatives. However, the above formulation is obviously a formal one and not a vigorous one.

Thus, we have to re-examine the function of one variable case to see if we could rephrase the concept of differentiability there in a new way, so that the new concept could be generalized or extended to the n-D setting.

Consider now $\quad f'(a) = \lim_{h\to 0} \dfrac{f(a+h)-f(a)}{h}$

We set $\delta(a,h) = \Delta y - dy$ as we go from $x=a$ to $x=a+h$ (see figure)



which is the error term or deviation when we approximate Δy by dy. Thus, we have

$$\boxed{\underbrace{f(a+h)-f(a)}_{\Delta y} = \underbrace{f'(a)\,h}_{dy} + \delta(a,h)} \quad \text{-------} \quad (**)$$

On dividing both sides by h, $\dfrac{f(a+h)-f(a)}{h} = f'(a) + \dfrac{\delta(a,h)}{h}$.

Since $f'(a) = \lim_{h\to 0} \dfrac{f(a+h)-f(a)}{h}$, we must have

$$\boxed{\lim_{h\to 0} \frac{\delta(a,h)}{h} = 0} \quad \text{-------} \quad (***)$$

Hence, while $\delta(a,h) \to 0$ at the same time as $h \to 0$, $\delta(a,h) \to 0$ much more faster than $h \to 0$. We'll express this subtle relationship between $\delta(a,h)$ and h in the new notation

$$\boxed{\delta(a,h) = o(h) \text{ as } h \to 0} \quad \text{-------} \quad (4*)$$

which is actually equivalent to $(***)$. Finally we could rewrite $(**)$ as

$$\boxed{f(a+h) - f(a) = f'(a)h + o(h) \text{ as } h \to 0} \quad \text{-------} \quad (5*)$$

Ex Consider $f(x) = x^2$ ∴ $f'(x) = 2x$, we have

$$f(a+h) - f(a) = f'(a)h + \delta(a,h)$$
$$\Rightarrow (a+h)^2 - a^2 = 2ah + \delta(a,h)$$
$$\Rightarrow \delta(a,h) = h^2.$$

As a result, $\dfrac{\delta(a,h)}{h} = h \to 0$ as $h \to 0$ ∴ $\delta(a,h) = o(h)$ //.

Ex Consider $f(x) = \sin x$, $f'(x) = \cos x$

$$f(a+h) - f(a) = \sin(a+h) - \sin a$$

Invoking the trigonometric identity $\sin\beta - \sin\theta = 2\cos\left(\frac{\beta+\theta}{2}\right)\sin\left(\frac{\beta-\theta}{2}\right)$,

$$f(a+h) - f(a) = 2\cos\left(a+\tfrac{h}{2}\right)\sin\left(\tfrac{h}{2}\right) = (\cos a)h + \delta(a,h)$$
$$\Rightarrow \delta(a,h) = 2\cos\left(a+\tfrac{h}{2}\right)\sin\left(\tfrac{h}{2}\right) - (\cos a)h$$
$$\Rightarrow \frac{\delta(a,h)}{h} = \cos\left(a+\tfrac{h}{2}\right)\frac{\sin(h/2)}{(h/2)} - \cos a$$
$$\Rightarrow \lim_{h\to 0} \frac{\delta(a,h)}{h} = 0 \quad \text{using the limit } \lim_{\theta\to 0}\left(\frac{\sin\theta}{\theta}\right) = 1$$

Now we are ready to put down a statement similar to (**) for $f(\vec{x})$, $\vec{x} = \langle x_1, \cdots, x_n\rangle$ in a n-D setting.

Indeed, set $\vec{a} = \langle a_1, \cdots, a_n\rangle$ and $\vec{h} = \langle h_1, \cdots, h_n\rangle$ which is a vector change in $\vec{x} = \langle x_1, \cdots, x_n\rangle$ as $\vec{x}$ changes from $\vec{a}$ to $\vec{a}+\vec{h}$, we want to put down

$$\underbrace{f(\vec{a}+\vec{h}) - f(\vec{a})}_{\text{scalar}} = \underbrace{? \cdot \vec{h}}_{\substack{\text{need to}\\ \text{be a scalar too}}} + \underbrace{\delta(\vec{a},\vec{h})}_{\text{scalar}}$$

? needs to be a vector $\vec{T} = \langle T_1, \cdots, T_n\rangle \in V_n$

where $\delta(\vec{a},\vec{h})$ is the error term when we approximate the change in f, $f(\vec{a}+\vec{h})-f(\vec{a})$ by $\vec{T}\cdot\vec{h}$. As $\vec{h}\to\vec{0}$, $\delta(\vec{a},\vec{h})\to 0$ far more rapidly than $|\vec{h}|\to 0$ i.e. $\lim_{\vec{h}\to\vec{0}} \frac{\delta(\vec{a},\vec{h})}{|\vec{h}|}=0$.

This motivates the following definition of Total Derivative.

Def^n. Consider a multi-variable function $f(\vec{x})$ with $\vec{x}=\langle x_1,\cdots,x_n\rangle$, let $\vec{a}=\langle a_1,\cdots,a_n\rangle\in D(f)$, the domain of f. Then f is said to be differentiable (or totally differentiable) at $\vec{a}$ iff $\exists\ \vec{T}=\langle T_1,\cdots,T_n\rangle$ such that

$$f(\vec{a}+\vec{h})-f(\vec{a})=\vec{T}\cdot\vec{h}+o(|\vec{h}|)$$

Here $o(|\vec{h}|)$ denotes a scalar error term which goes to zero more rapidly than $\vec{h}\to\vec{0}$ or $|\vec{h}|\to 0$.

$\vec{T}$ is known as the total derivative of f at the pt. $\vec{a}$.

Remarks:

(i) We might even denote $\vec{T}$ by $f'(\vec{a})$

(ii) The defining equation for $\vec{T}$ could also be rephrased as; $f(\vec{x})$ is said to be differentiable at the pt. $\vec{a}$ with total derivative $f'(\vec{a})=\vec{T}$ iff $\lim_{\vec{h}\to\vec{0}} \frac{f(\vec{a}+\vec{h})-f(\vec{a})-\vec{T}\cdot\vec{h}}{|\vec{h}|}=0$.

This formulation is more in the form of " $f'(a)=\lim_{h\to 0}\frac{f(a+h)-f(a)}{h}$ ".

Th^m Given $f(\vec{x})$, $\vec{x}=\langle x_1,\cdots,x_n\rangle$, assuming the total derivative $f'(\vec{a})$ exists at $\vec{a}\in D(f)$, then the partial derivatives $\frac{\partial f}{\partial x_1}(\vec{a}),\cdots,\frac{\partial f}{\partial x_n}(\vec{a})$ all exist and $f'(\vec{a})=\left\langle \frac{\partial f}{\partial x_1}(\vec{a}),\cdots,\frac{\partial f}{\partial x_n}(\vec{a})\right\rangle$.

i.e. the total derivative of f, if exists at $\vec{a}$, must coincide with $\nabla f(\vec{a})$, the gradient vector of f at $\vec{a}$.

Pf: Assignment //.

Remarks:

(i) $f'(\vec{a}) = \nabla f(\vec{a})$ is not a scalar but a vector in V_n (space of all the vectors in $\mathbb{R}^n$). From the geometric view point, it is a linear transformation $T|\mathbb{R}^n \to \mathbb{R}^1$ mapping the vector change $\vec{h} = \langle h_1, \cdots, h_n \rangle$ in $\vec{x}$ into $\mathbb{R}^1$.

i.e. $T\vec{h} = \nabla f(\vec{a}) \cdot \vec{h} = \frac{\partial f}{\partial x_1}(\vec{a}) h_1 + \cdots + \frac{\partial f}{\partial x_n} h_n$ (note that we use T instead of $\vec{T}$ to emphasize that T is a linear transformation).

(ii) In vector notation, $\overbrace{f(\vec{a}+\vec{h}) - f(\vec{a})}^{\Delta y} = \underbrace{\nabla f(\vec{a}) \cdot \vec{h}}_{dy} + o(|\vec{h}|)$. Now that for $\vec{x}$ nearby $\vec{a}$, if we set $\vec{x} = \vec{a} + \vec{h}$ so that $\vec{h} = \vec{x} - \vec{a}$, we have

$$f(\vec{x}) - f(\vec{a}) = \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}) + o(|\vec{x} - \vec{a}|)$$

$$\Rightarrow f(\vec{x}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a}) + o(|\vec{x} - \vec{a}|)$$

$$\Rightarrow \boxed{f(\vec{x}) = L(\vec{x}) + o(|\vec{x} - \vec{a}|)}$$

where $L(\vec{x}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot (\vec{x} - \vec{a})$ is the linear approximation of f at the point $\vec{a}$ (or the equation of tangent plane to the surface $w = f(\vec{x})$ at the point $(\vec{a}, f(\vec{a}))$).

The error term when we approximate $f(\vec{x})$ by $L(\vec{x})$ at the pt. a is of order $o(|\vec{x} - \vec{a}|)$ showing $L(\vec{x})$ is a good approximation because the error term $\to 0$ much more rapidly than $\vec{x} \to \vec{a}$ (or $|\vec{x} - \vec{a}| \to 0$)

Ex. $f(x,y) = x^2 + y^2$, show and find $f'(1,3)$.

Solution:

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = 2y$$

$f'(1,3)$ if exists must be given by $\nabla f(1,3) = \langle 2, 6 \rangle$.

It remains to verify the "o" condition i.e.

$$f(\vec{x}) = L(\vec{x}) + o(|\vec{x} - \vec{a}|)$$

where $L(\vec{x}) = f(1,3) + \nabla f(1,3) \cdot \langle x-1, y-3 \rangle = 10 + \langle 2, 6 \rangle \cdot \langle x-1, y-3 \rangle$

$= 10 + 2(x-1) + 6(y-3) = 2x + 6y - 10$

Set $\delta(\vec{x}) = f(\vec{x}) - L(\vec{x}) = x^2 + y^2 - 2x - 6y + 10$

$$= (x^2 - 2x + 1) + (y^2 - 6y + 9)$$

$$= (x-1)^2 + (y-3)^2$$

$$\lim_{(x,y)\to(1,3)} \delta(x,y) = \lim_{(x,y)\to(1,3)} \frac{(x-1)^2 + (y-3)^2}{\sqrt{(x-1)^2 + (y-3)^2}} = \lim_{(x,y)\to(1,3)} \sqrt{(x-1)^2 + (y-3)^2} = 0$$

Hence, $\vec{f}'(1,3) = \langle 2, 6 \rangle$ //.

Th^m. (Fundamental Theorem of Differential Calculus)

Given $f(\vec{x})$, $\vec{x} \in D$, D being the domain of f in $\mathbb{R}^n$. Suppose $\frac{\partial f}{\partial x_1}(\vec{a}), \dots, \frac{\partial f}{\partial x_n}(\vec{a})$ exists and are continuous $\forall a \in D$; then we have,

$$f(\vec{a} + \vec{h}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot \vec{h} + \vec{\varepsilon}(\vec{h}) \cdot \vec{h}$$

where $\vec{\varepsilon}(\vec{h}) = \langle \varepsilon_1(\vec{h}), \dots, \varepsilon_n(\vec{h}) \rangle$ satisfies

$$\lim_{|\vec{h}|\to 0} \varepsilon_1(\vec{h}) = \dots = \lim_{|\vec{h}|\to 0} \varepsilon_n(\vec{h}) = 0$$

Remarks:

(i) $\vec{\varepsilon}(\vec{h}) \cdot \vec{h}$ is the error term when we approximate $\Delta f = f(\vec{a} + \vec{h}) - f(\vec{a})$ by $df = \nabla f(\vec{a}) \cdot \vec{h}$.

(ii) Since $\frac{|\vec{\varepsilon}(\vec{h}) \cdot \vec{h}|}{|\vec{h}|} \leq |\vec{\varepsilon}(\vec{h})| = \sqrt{\varepsilon_1(\vec{h})^2 + \dots + \varepsilon_n(\vec{h})^2}$.

$$\lim_{|\vec{h}|\to 0} \frac{|\vec{\varepsilon}(\vec{h}) \cdot \vec{h}|}{|\vec{h}|} = 0 \Rightarrow f(\vec{a} + \vec{h}) = f(\vec{a}) + \nabla f(\vec{a}) \cdot \vec{h} + o(|\vec{h}|)$$

implying f is differentiable at $\vec{a}$ with $f'(\vec{a}) = \nabla f(\vec{a})$.

Hence, if f has all its 1st order partial derivatives exist and are continuous, then f is also differentiable with its total derivative $f' = \nabla f$.